Evidence for Electroweak Single Top Quark Production at D0

Leonard Christofek
Department of Physics, Brown University, Box 1843,
Providence, RI 02912

We present first evidence for electroweak single top quark production from the D0 Collaboration at the Fermilab Tevatron Collider. Using a 0.9 fb^{-1} data sample, several multivariate techniques are used to isolate the single top signal from background. Combining these three methods (Bayesian Neural Networks, Decision Trees, and Matrix Elements), we obtain a measured cross section of $\sigma(p\overline{p} \rightarrow tb + X, tqb + X) = 4.8 \pm 1.3$ pb with a Gaussian significance of 3.5σ. Using this measurement, we set a lower 95% C.L. on the V_{tb} element of the Cabibbo-Kobayashi-Maskawa (CKM) matrix of $0.68 < |V_{tb}| \leq 1$.

1 The Single Top Analysis

1.1 Theory and Motivation

Examining single top quark production at the Tevatron [1], we find two major production modes for single top and Figure 1 shows the two representative Feynman diagrams. The first one is called s-channel production where two quarks annihilate to make an off-shell W boson which then decays to a top quark and a bottom quark. This process is also known as tb because it has a top and bottom quark in the final state. The second process, or t-channel production, is when a light quark emits a W boson which fuses with a b-quark coming from a gluon that splits into a $b\overline{b}$ pair to make a top quark. Because this process has a top, a bottom, and a light quark in the final state, it is also known as tqb. [a]

Observation of single top quark production allows one to study the Vtb coupling and directly measure the CKM element V_{tb} without the standard model (SM) assumption of three families. In addition, the single top cross section is also sensitive to new physics. Since single top is

[a] There is a third process called associated production, but the cross section is too small to observe at the Tevatron, so we will not discuss it further.

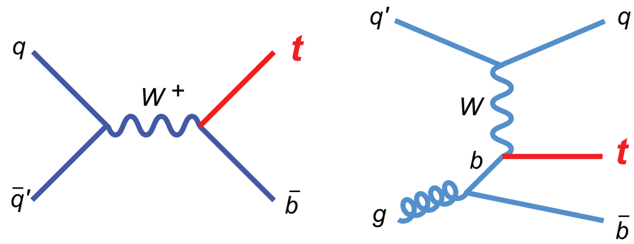

Figure 1: Representative Feynman diagrams for single top quark production. The left plot show the tb production mode and the right plot shows the tqb production mode.

a background for the Higgs search, this analysis is a "proof of principle" for the advanced techniques that are being used.

1.2 Experimental Signature, Backgrounds, and Data Samples

The experimental signature for single top is a high p_T lepton, large missing transverse energy corresponding to the neutrino, and two to four jets (this is done to maximize acceptance). One can further enhance the signal content of the sample by requiring that either one or two jets are associated with b-quarks (also known as b-tagging).

There are three main backgrounds to single top. One comes from $t\bar{t}$ pair production which is normalized to the NNLO cross section. Then there is multijet production where a jet fluctuates to mimic a lepton which also produces missing energy. And W+jet production which includes $Wb\bar{b}, Wc\bar{c}$, and Wjj. The multijet and W+jet background yields are normalized to data before b-tagging.

We take our full data sample and divide it into electrons and muons. These samples are further separated by the number of jets in the event. Finally, we divide these samples again by determining how many jets can be associated with b-quark(s). The event yields and systematic uncertainties are shown in Table 1.2[b]. The acceptance for the single top signal is $(3.2 \pm 0.4)\%$ for tb and $(2.1 \pm 0.3)\%$ for tqb. In addition, we have checked the agreement between data and background model for over 100 individual distributions and find good agreement.

Source	2 jets	3 jets	4 jets
tb	16±3	8±2	2±1
tqb	20±4	12±3	4±1
$t\bar{t} \to \ell\ell$	39±9	32±7	11±3
$t\bar{t} \to \ell$+jets	20±5	103±25	143±33
$Wb\bar{b}$	261±55	120±24	35±7
$Wc\bar{c}$	151±31	85±17	23±5
Wjj	119±25	43±9	12±2
Multijets	95±19	77±15	29±6
Total bkgd	686±41	460±39	253±38
Data	697	455	246

Source	Size
Top pair normalization	18%
W+jet & multijet normalization	18-28%
Integrated luminosity	6%
Trigger modeling	3-6%
Lepton ID corrections	2-7%
Other small components	Few%
Jet energy calibration	1-20%
b-tagged jet modeling	2-16%

Table 1: The left table show the numbers of expected and observed events in 0.9 fb^{-1} for e and μ, 1 b tag and 2 b tag channels combined. The right table shows the sizes of the systematic uncertainties for this analysis.

[b] An important point to note is that the single top signal is smaller than the background uncertainty.

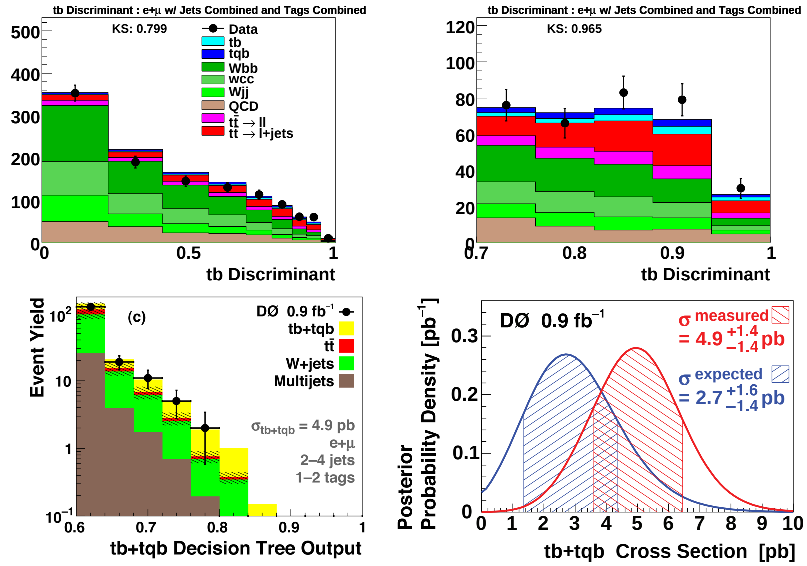

Figure 2: The top left plots shows the full discriminant for the Matrix Element Analysis. The top right plots shows an enlarged region of the discriminant near the value one. The bottom left plot shows an enlarged region of the Decision Tree output for high values of the discriminant. The bottom right plot shows the posterior distributions for the expected and measured cross sections for the Decision Tree analysis.

	Bayesian NN		Matrix Element		Decision Trees	
	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.
$\sigma(p\overline{p} \to tb + X, tqb + X)$ [pb]	$3.2^{+2.0}_{-1.8}$	5.0 ± 1.9	$3.0^{+1.8}_{-1.5}$	$4.6^{+1.8}_{-1.5}$	$2.7^{+1.6}_{-1.4}$	4.9 ± 1.4
Significance	1.3σ	2.4σ	1.8σ	2.9σ	2.1σ	3.4σ

Table 2: Table of expected and observed cross sections along with the expected sensitivities and observed significances.

1.3 Multivariate Methods and Results

We have used three methods to separate the single top signal from the background. The first is called Bayesian Neural Networks [2], the second is called Decision Trees [3], and the third is called Matrix Elements. Then using the outputs from each method we construct a binned likelihood and extract a cross section.

The first method uses Bayesian Neural Networks. It uses 24 input variables for training the networks (where the signal tends towards one and the background towards zero). A simple description of BNNs is that of averaging many individual NNs. The second method uses Decision Trees. This analysis uses 49 input variables, and the idea is to recover events that are rejected by a simple cut-based analysis. In addition, for events that get misclassified, we use boosting which effectively averages over many trees to improve signal and background separation. The third method used is called the Matrix Element Method. This is a different idea based on using the full kinematic information from the reconstructed objects in the event to form a discriminant to separate signal from background.

In Figure 2, we show the discriminant output for the Matrix Element and Decision Tree analyses along with the expected and measured cross sections from the Decision Tree analysis. In Table 1.3, we show the expected sensitivities and the measured values of the cross sections along with the significances. Decision trees have a 3.4σ excess which establishes the first evidence for single top quark production [4].

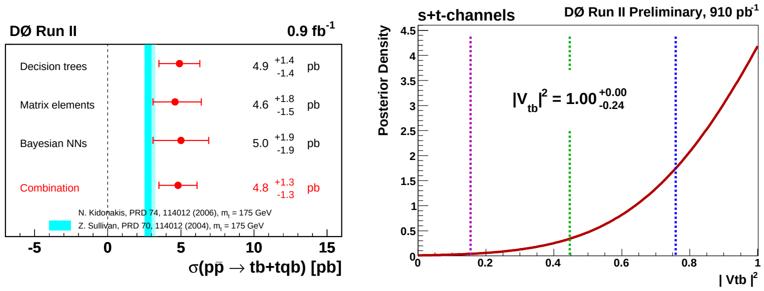

Figure 3: The left plot shows the combination of the three techniques used in this analysis. The right plot shows the lower limits obtained on $|V_{tb}|$.

The most general tbW vertex is given by [5]:

$$\Gamma^{\mu}_{tbW} = -\frac{g}{\sqrt{2}} V_{tb} \bar{u}(p_b) \left[\gamma^{\mu} (f_1^L P_L + f_1^R P_R) - \frac{i\sigma^{\mu\nu}}{M_W} (f_2^L P_L + f_2^R P_R) \right] u(p_t), \qquad (1)$$

where k is the W four-momentum and the f_1 and f_2 couplings can a-priori be CP-violating. In the case of the SM, CP is conserved in the tbW vertex and $f_1^L = 1$ and $f_1^R = f_2^L = f_2^R = 0$. Since Decision Trees are the most sensitive result, with a measured cross section of $\sigma(p\overline{p} \rightarrow tb + X, tqb + X) = 4.9 \pm 1.4$ pb, we use this value to extract a value of $|V_{tb} f_1^L| = 1.3 \pm 0.2$ and assuming full SM values (i.e. $f_1^L = 1$) we set a lower limit on $|V_{tb}|$ at 95% C.L. of $0.68 < |V_{tb}| \leq 1$. We have combined the outputs of these three measurements using a simple linear method and find a combined cross section of 4.8 ± 1.3 pb with a significance of $3.5\ \sigma$. We show these results in Figure 3.

In summary, we have found first evidence for single top quark production with a measured cross section $\sigma(p\overline{p} \rightarrow tb + X, tqb + X) = 4.8 \pm 1.3$ pb. Using this measurement we have set lower limits on $|V_{tb}|$ of $0.68 < |V_{tb}| \leq 1$.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); PPARC (United Kingdom); MSMT (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); Research Corporation; Alexander von Humboldt Foundation; and the Marie Curie Program.

References

1. Z. Sullivan, Phys. Rev. D **70**, 114012 (2004).
2. R.M. Neal, *Bayesian Learning of Neural Networks*, (Springer-Verlag, New York, 1996).
3. L. Brieman *et al., Classification and Regression Trees* Wadsworth, Stamford, 1984); D. Bowser-Chao and D.L. Dzialo, Phys.Rev. D **47**, 1900(1993).
4. V.M. Abazov, *et al*, Phys. Rev. Lett. bf 98, 181802 (2007).
5. G.L. Kane, F.A. Ladinsky and C.P. Yuan, Phys. Rev. D **45**, 124 (1992); K. Whisnant *et al.*, Phys. Rev. D **56**, 467 (1997).